```
   1   AACAATTGCC GCGAATTCGG CACGAGATGA AATCTAGTTG TTTAAAAGCG
  51   TGTAGCACCT CCTCCCTCTC TCTTACTCCT GCTCTCACCA TGTGAGACGC
 101   CTCGCTCCCC CTTTGCCTTT CACCAGGATT GGAAGCTTCC TGAGGCCTCC
 151   CCAGAAGCAG AAGCTGCTAT GCTTCTTGTA CAGTCTGTAG AGCTATTAGC
 201   CAGTTAAACC CATTTCCTTC ATAAATTTCC CAGTCTCAGG TATTTCTTTT
 251   TAGCAATTTG AGAATGAACT AATACACAGA CAGAGAGCCA GGAGATGGAA
 301   ATCCCAAGGT GCTTTCCTGC TGTCTTCCAG TCTCCTGCTG GTGTCTCCCA
 351   GTGTCTCAAT TCCACCAGAA ACCAGAAATA AAAAGAATCC CACTGATGTG
 401   GTACATAGAA GCCACTCTCT TGGGATGTCA AACAGGATAA AGAAGAATGG
 451   AAAGCAAATC CTCATGGGTC ATCAGACTGG GGTTTCTGAG CATGGATTCA
 501   ACCATCCCAG TCTTGGGTAC AGAACTGACA CCAATCAACG GACGTGAGGA
 551   GACTCCTTGC TACAAGCAGA CCCTGAGCTT CACGGGGCTG ACGTGCATCG
 601   TTTCCCTTGT CGCGCTGACA GGAAACGCGG TTGTGCTCTG GCTCCTGGGC
 651   TGCCGCATGC GCAGGAACGC TGTCTCCATC TACATCCTCA ACCTGGTCGC
 701   GGCCGACTTC CTCTTCCTTA GCGGCCACAT TATATGTTCG CCGTTACGCC
 751   TCATCAATAT CCGCCATCCC ATCTCCAAAA TCCTCAGTCC TGTGATGACC
 801   TTTCCCTACT TTATAGGCCT AAGCATGCTG AGCGCCATCA GCACCGAGCG
 851   CTGCCTGTCC ATCCTGTGGC CCATCTGGTA CCACTGCCGC CGCCCCAGAT
 901   ACCTGTCATC GGTCATGTGT GTCCTGCTCT GGGCCCTGTC CCTGCTGCGG
 951   AGTATCCTGG AGTGGATGTT CTGTGACTTC CTGTTTAGTG GTGCTGATTC
1001   TGTTTGGTGT GAAACGTCAG ATTTCATTAC AATCGCGTGG CTGGTTTTTT
1051   TATGTGTGGT TCTCTGTGGG TCCAGCCTGG TCCTGCTGGT CAGGATTCTC
1101   TGTGGATCCC GGAAGATGCC GCTGACCAGG CTGTACGTGA CCATCCTCCT
1151   CACAGTGCTG GTCTTCCTCC TCTGTGGCCT GCCCTTTGGC ATTCAGTGGG
1201   CCCTGTTTTC CAGGATCCAC CTGGATTGGA AAGTCTTATT TTGTCATGTG
1251   CATCTAGTTT CCATTTTCCT GTCCGCTCTT AACAGCAGTG CCAACCCCAT
1301   CATTTACTTC TTCGTGGGCT CCTTTAGGCA GCGTCAAAAT AGGCAGAACC
1351   TGAAGCTGGT TCTCCAGAGG GCTCTGCAGG ACACGCCTGA GGTGGATGAA
1401   GGTGGAGGGT GGCTTCCTCA GGAAACCCTG GAGCTGTCGG GAAGCAGATT
1451   GGAGCAGTGA GGAAGAACCT CTGCCCTGTC AGACAGGACT TTGAGAGCAA
1501   TGCTGCCCTG CCACCCTTGA CAATTATATG CATTTTTCTT AGCCTTCTGC
1551   CTCAGAAATG TCTCAGTGGT CCCTCAAGGT CTTCGAATAG ATGTTTATCT
1601   AACCTGACAG TTGCAGTTTT CACCCATGGA AAGCATTAGT CTGACAGTAC
1651   AATGTTTGGA TTCTCCTTGA TATTACCAAT ACATTTTCCC TGTTATCTTG
1701   CACTGAATCT TTCCTACTGA ACACTTTTTC TGCACTTTTC ATTGTAATAA
1751   AAGGAGTTGC TGTCCACAAC CCTAAAACTC TTCTTTATAC TTGTTTCCTA
1801   CCTGATAGTA TCAAAAAGGA AGATTCCTTA TTAATCTGTC AGACTATGTT
1851   CCCCTGAAAA TCATGTTCCC TTTTATGACT GGAGGCATTA CTGCAGTTGG
1901   AAGCTCAATT CTTAATAAGT GAGTTCTGCT ACCTCTAAAT TCCATTGAAT
1951   TCTCAGATAT AAAGCAAAAT AATGACCTTA GAGAGAGATT CTCCCTTCAT
2001   AAAAACAGTC TTAGAAATTG GTTTTATGAA TAGCCCTCTC CTGTCATTTG
2051   TCCACAGCAT GGTGACATGT TGGCCTTGGT TTCTAGTAAA GACAATCGTG
2101   GCCCCTTCCC CTTGAGAACT GGTAAGTTCT TATTTAGCTC TTCCTGGACT
2151   AATGAACTAG TGAGGAGCCT ATAAATATGT CCCACCAGTT TCATTTTGGC
2201   CATTGGAAAC CTCAATATTG ATTTTAAAGT GGAAATTATC TTGAAAACCA
2251   TTTATTATTC ACTTACAGAT TCTTTCAGTT GTAGGAGAAT TCTTCATACT
2301   TCCAGGTTTT GTATAAATTG TTCTGATTGT AACTTTCAGT TAGTTTTTATG
2351   GCTGTTTACA TGAGAAGCAA AACTGAAAAC ATCTGACCTT TCCATGACAA
2401   TCTCAATTAT GGTATCTGGA TAATAACTTA CAGTTGGTAC AGAATTCTGA
2451   TACATGCTGT GACATACATG AACCTGGAAA TATTGTGCTA AGGAAAATAA
2501   GCCAGACGCC AAACAATATT GTAAGTTCAA ATTCTATGAG GTATCCAAAT
2551   TAGGAAATTC TTGAACACAG AAAATAAATT AGGAGGATCC TGGTGCTGGA
2601   AAAAAAAAAA AAAAAAA   (SEQ ID NO:1)
```

FEATURES:
Start: 447
Stop: 1458

HOMOLOGOUS PROTEIN:
Top BLAST Hits:

	Score	E
gi|547920|sp|P35410|MRG_HUMAN MAS-RELATED G PROTEIN-COUPLED REC...	174	1e-42

FIGURE 1, page 1 of 2

```
gi|6981186|ref|NP_036889.1| MAS1 oncogene >gi|135921|sp|P12526|...    170   2e-41
gi|4505105|ref|NP_002368.1| MAS1 oncogene >gi|135920|sp|P04201|...    167   2e-40
gi|6678804|ref|NP_032578.1| MAS1 oncogene >gi|266505|sp|P30554|...    163   3e-39
gi|2118485|pir||S51001 transforming protein mas - mouse              163   3e-39
gi|134079|sp|P23749|RTA_RAT PROBABLE G PROTEIN-COUPLED RECEPTOR...    142   6e-33
gi|4455061|gb|AAD21055.1| (AF118265) orphan G protein-coupled r...     89   7e-17
gi|4758070|ref|NP_004769.1| G protein-coupled receptor 44 >gi|4...     89   7e-17
gi|3023772|sp|P79243|FML2_PANTR N-FORMYL PEPTIDE RECEPTOR-LIKE ...     84   2e-15
gi|6753528|ref|NP_034092.1| chemoattractant receptor-homologous...     83   3e-15
gi|3023793|sp|P79237|FML2_PONPY N-FORMYL PEPTIDE RECEPTOR-LIKE ...     83   5e-15
gi|292035|gb|AAA52474.1| (L14061) N-formyl peptide receptor-lik...     82   9e-15
gi|3023767|sp|P79178|FML2_GORGO FMLP-RELATED RECEPTOR II (FMLP-...     82   9e-15
```

BLAST dbEST hit:

```
                                                                    Score     E
gi|2253096|gb|AF003828.1|AF003828 AF003828 Human erythroleukemi...   165   4e-38
```

EXPRESSION INFORMATION FOR MODULATORY USE:
Expression information from BLAST dbEST hit:
```
gi|2253096|gb|AF003828.1   Human erythroleukemia
```

Tissue expression from PCR-based tissue screening panels:
Human testis

FIGURE 1, page 2 of 2

```
  1  MESKSSWVIR LGFLSMDSTI PVLGTELTPI NGREETPCYK QTLSFTGLTC
 51  IVSLVALTGN AVVLWLLGCR MRRNAVSIYI LNLVAADFLF LSGHIICSPL
101  RLINIRHPIS KILSPVMTFP YFIGLSMLSA ISTERCLSIL WPIWYHCRRP
151  RYLSSVMCVL LWALSLLRSI LEWMFCDFLF SGADSVWCET SDFITIAWLV
201  FLCVVLCGSS LVLLVRILCG SRKMPLTRLY VTILLTVLVF LLCGLPFGIQ
251  WALFSRIHLD WKVLFCHVHL VSIFLSALNS SANPIIYFFV GSFRQRQNRQ
301  NLKLVLQRAL QDTPEVDEGG GWLPQETLEL SGSRLEQ  (SEQ ID NO:2)
```

FEATURES:

Functional domains and key regions:

[1] PDOC00001 PS00001 ASN_GLYCOSYLATIONN-glycosylation site
         279-282 NSSA
_____
[2] PDOC00005 PS00005 PKC_PHOSPHO_SITEProtein kinase C phosphorylation site
Number of matches: 3
     1   133-135 TER
     2   221-223 SRK
     3   292-294 SFR
_____
[3] PDOC00006 PS00006 CK2_PHOSPHO_SITECasein kinase II phosphorylation site
Number of matches: 3
     1   169-172 SILE
     2   181-184 SGAD
     3   333-336 SRLE
_____
[4] PDOC00008 PS00008 MYRISTYLN-myristoylation site
Number of matches: 2
     1   244-249 GLPFGI
     2   248-253 GIQWAL
_____
[5] PDOC00210 PS00237 G_PROTEIN_RECEP_F1_1G-protein coupled receptors family 1
signature

Membrane spanning structure and domains:

Helix	Begin	End	Score	Certainty
1	41	61	1.775	Certain
2	75	95	1.059	Certain
3	112	132	1.947	Certain
4	151	171	1.380	Certain
5	193	213	2.255	Certain
6	229	249	2.322	Certain
7	261	281	1.221	Certain

FIGURE 2, page 1 of 2

BLAST Alignment to Top Hit:
>gi|547920|sp|P35410|MRG_HUMAN MAS-RELATED G PROTEIN-COUPLED
         RECEPTOR MRG >gi|320141|pir||A39485 transforming protein
         (mrg) - human >gi|244210|gb|AAB21255.1| (S78653) mas
         product homolog modulating intracellular angiotensin II
         actions=mrg [human, Peptide, 378 aa] [Homo sapiens]
         Length = 378

 Score = 174 bits (437), Expect = 1e-42
 Identities = 104/275 (37%), Positives = 161/275 (57%), Gaps = 24/275 (8%)

Query: 51 IVSLVALTGNAVVLWLLGCRMRRNAVSIYILNLVAADFLFLS----GHIICSPLRLINIR 106
           +VSL + N V WLL C   N   +YIL+LVAAD ++L   G + + L   +
Sbjct: 84 LVSLCGVLLNGTVFWLLCCGAT-NPYMVYILHLVAADVIYLCCSAVGFLQVTLLTYHGVV 142

Query: 107 HPISKILSPVMTFPYFIGLSMLSAISTERCLSILWPIWYHCRRPRYLSSVMCVLLWALSL 166
           I   L+ + F + + +L AISTERC+ +L+PIWY C RP+Y S+V+C L+W L
Sbjct: 143 FFIPDFLAILSPFSFEVCLCLLVAISTERCVCVLFPIWYRCHRPKYTSNVVCTLIWGLPF 202

Query: 167 LRSILEWMFCDFLFSGADSVWCETSD---FITIAWL--VFLCVVLCGSSLVLLVRILCGS 221
           +I++ +F +   W   F+ ++ L   L +V+C SSL LL+R LC S
Sbjct: 203 CINIVKSLFLTY--------WKHVKACVIFLKLSGLFHAILSLVMCVSSLTLLIRFLCCS 254

Query: 222 RKMPLTRLYVTILLTVLVFLLCGLPFGIQWALFSRIHLDWKVLFCHVHLVSIFLSALNSS 281
           ++   TR+Y + ++ +FLL LP +   + + D+K+   +L+S+FL +NSS
Sbjct: 255 QQQKATRVYAVVQISAPMFLLWALPLSV-----APLITDFKMFVTTSYLISLFL-IINSS 308

Query: 282 ANPIIYFFVGSFRQRQNRQNLKLVLQRALQDTPEV 316
           ANPIIYFFVGS R+++ +++L+++LQRAL D PEV
Sbjct: 309 ANPIIYFFVGSLRKKRLKESLRVILQRALADKPEV 343   (SEQ ID NO:4)

Hmmer search results (Pfam):
Scores for sequence family classification (score includes all domains):

Model	Description	Score	E-value	N
PF00131	Metallothionein	382.6	3.9e-111	18
PF00956	Nucleosome assembly protein (NAP)	27.6	9.6e-07	3
CE00408	E00408 osteopontin	26.2	2e-06	3
PF00183	Hsp90 protein	24.0	2.8e-05	3
PF00037	4Fe-4S ferredoxins and related iron-sulfur c	20.9	7.2e-05	6
PF01056	Myc amino-terminal region	19.5	6.3e-06	3
PF00524	E1 Protein, N terminal domain	16.4	0.00089	4
PF01448	ELM2 domain	13.5	0.012	3
PF00428	60s Acidic ribosomal protein	12.5	0.0062	3
PF00095	WAP-type (Whey Acidic Protein) 'four-disulfi	11.2	0.23	2
PF01025	GrpE	8.2	0.28	2
PF01437	Plexin repeat	6.4	1	3
PF00057	Low-density lipoprotein receptor domain clas	6.2	2.7	6
PF00007	Cystine-knot domain	5.9	1.4	5
CE00299	CE00299 fibromodulin	5.2	1.3	2
PF00020	TNFR/NGFR cysteine-rich region	4.3	8.9	1
PF01258	Prokaryotic dksA/traR C4-type zinc finger	4.3	7	1
PF00865	Osteopontin	2.7	3.3	1
PF00913	Trypanosome variant surface glycoprotein	2.2	8.2	1
CE00545	CE00545 progesteron_receptor	1.7	1.8	2
CE00412	E00412 BRCA1	1.7	5.1	1
PF01216	Calsequestrin	0.8	9.2	1
CE00038	CE00038 calcium_channel_L_type	-0.1	3.5	1

FIGURE 2, page 2 of 2

```
   1  TGTATGAAGC CAATGTCACT TTAATACCAA AACCAGGAAA GGATATACAA
  51  AAAAGAAAAC TATAGACCAG TACCACTGAT GAATATACAT GCAGAATACC
 101  CCAACAAAAT ACTAGCTAAC CCAATCCAAC AGCATATCAA GAAGATAATC
 151  CACCATTGTC AAGTGGGTTT CATCCCAGGG GTGCAGGATA GGTTAACATA
 201  CACAAGTCAA TAAATGTGAT ACATCACATA AACAGAATTA AAAACAAAAA
 251  TCACATGATC ATCTCAATAG ATGCTGAAAA AGCATTTGAC AAAATCTAAC
 301  ATTTCTTTAT GATTAAAACC TTCAGCAAAA TCGACATAGA AAGGACATAC
 351  CTTAATGTAA TAAAAGCCAT ATATGACGGA CCCACAGCAA ACATTATACT
 401  GAATGGGGAA AAGTTGAAAA CATTGTCCCT GAGAACTGGA ACAAGACAAG
 451  GATGCTACTT TCACCACTTC TATTCAACAT AGTAGTGGAA GTTTTAGCCA
 501  GAGCAATCAG ACAAGAGAAA GAAATCAAGG GCACCCAAAT CAATAAAGAG
 551  GAAGTCAAAC TGTCCCTGTT CACTGATGAT ATGATTGTAT ACCTAGAAAA
 601  CCCTAAAGAC TCATCCAGAA AGCTCCTAGA ACTGATACAT AAATTCAGTA
 651  AAGTTTCAGG ATACAAACTA AATGTACACA AATCAGTAGC ACTGCTATAC
 701  ACCAACAGTG ACCAAGCTGA GAATCAAATC AAGAACTCAA ACACTTTTAC
 751  AATAGCTGTA AAAAAATACT TAAGAATATT CTTACCCAAG GAGGTGAAGG
 801  ACCTCTACAA GGAAAACTAC AAAACACAGC TGACATCATA GATGACACAA
 851  ACAAGTGGAA ACACATCCCA TGCTCATGGA TGGGTAGAAT CAATATTGTG
 901  AAAATGACCA TATTGCCAAA AGCAATCTAC AAGTTCAATG CAATTCCCAC
 951  CAAAATATCA TCATCATTCT TCACAGAACT AGAAAAAAAC AATTCTAAAA
1001  TTCATATGGA ACAACAACCA AAAAAAAAAA AAAAAACCCG CATAGCCAAA
1051  GCAAGACTTA GCAAAAAGAA CAAATCTGGA GGCATCACAT TACCCATCTT
1101  CAAACTATAC TACAAGGCTA TAATCACCAA AACATCATGG CACTGACATA
1151  AAACTAGGCA CATAGACCAA TGGAAAAGAA GAGAGAATCC AGAAATAAAG
1201  CCAAATAATT ATAGCCAACT GATTTTTGAC AAAGCAAACA AAAACATAAA
1251  GTGGGGAAAA GACATTCTAG TTAACAAATG GTGCTGAGAT TATTGGCAAG
1301  CCACATGTGG AAGAATGAAA CTGGATCCCT TGTCTCTCAC TTAATACAAA
1351  AATTGATACA AGATGGATCA AAGACTTAAA TCTGAGACCT AAAACCATAA
1401  AAATTCTAGA AGATAACATC AGAAAAATGC TTCTAGACAT TCACTTAGGC
1451  AAAGACTTCA TGGCCAAGAA CCCAAAAGTA AATGCAACAA AAACAAAAAT
1501  AAATAGATAG GACTTAATTA AACTAAAAAG CTTTTGCGCA GCAAAAACAA
1551  TCATTAGCAG AGCAAACAGA CAACCCACCG AGTGAGAGAA AATCTTCACA
1601  AACTAAGCAT CTGACTAAGG ACTAATATCC GGAATCCACA AGGAACTCAA
1651  ACAAATCAGC AAGAAGAAAG CAAACAATCC CATGAAAGAG TGGGCTAAGG
1701  ACATGAATAG ACAATTCTCA AAAGAAGATA TACAAATGGC CAACAAACAG
1751  GAAAAAATGC TTAACATCAC TAATGATTAG GGAAATGTAA ATCAACACTG
1801  TAATGCGATA CCACCTTACT CCTGCAAGAA TGGTCATAAT TTAAAAATCT
1851  AAAAATAATA GATGTTGGTG GGTCTGTGGT GATAAAGGAA CACTTTTACA
1901  CTGCTGGTGG GAATGTAAAC TTGCGCAACC ACTATGGAAA ACAGTGTGGA
1951  AATTTCTTAA GGAACTAAAA GTAGATCGAC CATTTGATCC AGCAATCCCA
2001  TTAAATATGT ATAAATATAT ATATTTATAT ACCATGGAAT ACAACTCAGC
2051  CATAAAAAAG AATAAAATGA TGACATTCAC AGCAATCTAG ATGGAATTGG
2101  AGACCCTTAT TCTAAGTGGG GTAACTCAGG AATGGAAAAC CAAACATCAT
2151  ATGTTCTCAC TTACAAGTGG GGGCTAAGCT GTGAGGACAC GAAGGCATAG
2201  AATGATATAA TGAACTCTGG GGACTTGAGG GGAAGGATGG AAGAGAGGCG
2251  AGGGATAAAA GACTACACAA TGGGTACAGT GTACACTGCT CAGGTGATGG
2301  GTGCACCAAA ATCTCAGAAA TTACCACTAA AGAACTTATC CATGGAAGCA
2351  AACACCACCT GTTCCCCAAA ATCCCAATGA AATAAAAATA ATAATAATAA
2401  ATGATTTAAT TTCACAGAAT TTAAAAAAGT TCACTGTTCA GAGTTTATAA
2451  TAATGAAGTA AGAATGAAAA GTGTAGCAAG TGGTAGCCTC TGGACAATGG
2501  GACTCTAGAT TTTCACCTTG CATACACTTC TCTGGCATTT GGAAAGAAAG
2551  TATACACATG AATATATCAC CACTATGATA AAGAAAACAT CAAAAAATTG
2601  TGTCAGGCCA TTGTCAGCCT TGAATGGTCC CATGATCTAC TTTTTCATTT
2651  GGATATAAAG CCTCATAATG ATAGTTCACA TTGCTTAATG TGATGCCTAG
2701  GCCCATAATT GATTTTTAAA ATCAGGACAG CAATTACTTA CAGGAAGTTG
2751  AACAAGATGG GACGTGATAG GAGAGGCTTA AATGTACTGG ATATGGGACA
2801  GAGGCCAAGA ATCATCTCAG TTAGGATTTG TGTCTCAAAT ACCTCTGGCC
2851  TCTGATTTGC CCATAGTCCT CATACAGGAA ATAACAAGAC TGTCCAGCAT
2901  CTTCGTAAGC CTGGATTGCT CACCAGCTTT CATTTCAGCT CCTGTAGGCA
2951  TCTCCTGAAT TAAGCAACAC AGAAAAGTCC TCTGAAGTCA CTGAATCCCA
3001  GAAAGGCTCT CTACCTTTAG CACAAGGGAG GTCTTCACCA CTGGACAAAG
3051  AAGGAACGAT AAGGGTAAGT ACCAAGAACT CTCTTCTTCC ACAGTCAGTT
3101  ATGATTTTTG CTGTAAGATC ATGTCCTTAT GCTTCCACCT TGGTGCTACA
```

FIGURE 3, page 1 of 3

```
3151   TGCAGGGGGT CACGAGCCTG TTTCAGGAAA AGACAGGAGA CATGAAGCTT
3201   CCTTTCAGAA ACTGAGTGCT GTCAACCCAA ACTGTGTGAG CTCTAAATGG
3251   TGTCCCCCCT TCTAATTTAT CTCCCCATAT CACCTCCTTC ATTCCAATCA
3301   TTCAATCTGC CCTCATGGAG AGACTGCTGC CTCTTACATT CATTTAACGA
3351   GCAAGGGGAC ATGCAGGCAT TTCTTCCCAG AGTTGAACTG CTATAGAGCC
3401   AGTTTCTTTG TTTCACTTAC TTTTCAAATT TATTCTTCTT TGCCTATCTG
3451   GAAAGGTCTA AGGAAGATAT AGATGGCCCA ATAATTAAGG AGTGTTTCAT
3501   GAGGAAAGTA TTTACAAAGA TGCACAGAGT TAAGGGTCAG GATCCTAAGC
3551   AGCAATACAT AGGGGAGCAC TACTTCCTCC CCTAGGCTGA AACGGACAGG
3601   GAAGGAGCAG TTACCATTGT CGCCATAGCC ATAGCTGTAG CCATAAGGGT
3651   GGGAGAGCAT GAGCAGGCAA GTGGAGAAGC CCTGCGTGGC CAACGCACAG
3701   CCACACAGGC TGATATAGTT TGGATCTGTG TTCCCACCAA AATCTCATGT
3751   TGATTGTAAT TTCCAATGTT GGAGGAAGGG CCTTGTGGGA GATGATTATT
3801   AGATCACGGG GATGGTTTTG CATGAATGTT TTAACACCAT CCCCCTTTGG
3851   TATTGTTGTT GTGATACTGA CGAGTTCTCA TGAAATCTAG TTGTTTAAAA
3901   GCGTGTAGCA CCTCCTCCCT CTCTCTTACT CCTGCTCTCA CCATGTGAGA
3951   CGCCTCGCTC CCCCTTTGCC TTTCACCAGG ATTGGAAGCT TCCTGAGGCC
4001   TCCCCAGAAG CAGAAGCTGC TATGCTTCTT GTACAGTCTG TAGAGCTATT
4051   AGCCAGTTAA ACCCATTTCC TTCATAAATT TCCCAGTCTC AGGTATTTCT
4101   TTTTAGCAAT TTGAGAATGA ACTAATACAC AGACAGAGAG CCAGGAGATG
4151   GAAATCCCAA GGTGCTTTCC TGCTGTCTTC CAGTCTCCTG CTGGTGTCTC
4201   CCAGTGTCTC AATTCCACCA GAAACCAGAA ATAAAAAGAA TCCCACTGAT
4251   GTGGTACATA GAAGCCACTC TCTTGGGATG TCAAACAGGA TAAAGAAGAA
4301   TGGAAAGCAA ATCCTCATGG TAAATGAGAC TATCCCTCTC ACCTTCTTGT
4351   ATCCTCCTAA TTCCTGGGGC TTTCTCTATC TGATTGATCC CTGTCTCATT
4401   TCAGCTCTAT CAGACTACTT TAATGTTTGG CTTGTCTTTC TCTACTGTCA
4451   CTTTTATGCA GAAATGTTTG CATTTGTTAA AAATGCATAG AAAATAAAAT
4501   GTAATTTTAA AAAGAACATA TGTATTTTGT TTAGAATATA AGTTTGGCTG
4551   ATCTAATAAA GACATGAAGA AGAAATATCT TAAACAAGAA AGTATAGTTG
4601   TGCCTCTGGG TCACTAGGTT CTGAATCTAC AGATTCAACA AACTACAGGA
4651   GGAAACTTTT CCAAAAATAA AGGTGTGGCG GAGTTGTGTA TGTACTGAAC
4701   AGGTACAAAC TTGTATTTCT TTGTCATTAT TTCTGAAAAA CTACAATATA
4751   ACAAGAACTT ATATAGCATT TGCATTTTGT CAGTTATTCT AAATAACTTT
4801   AAATGATTTA ATGTATCTGG GAGAAAGTGC ATAGAGTATA TACAAATACC
4851   ATATATAAGG AAATTGAGCA TCTGCAGATT TTGGTCTGTG CTGGGGTTCT
4901   GGAAAGAATC CCCTGTAAAT ACACAAAAAT GACACTCTTC GAGATCTGAA
4951   CTAGAAGCTC CAAAGCATCA TACATCAGAA TTCCAAAAAT TGCTGCTCCC
5001   CAGTTCCTAG AGAGTTGCCC TCATCCTTGT GATCCTACAT GGTTCCCAGC
5051   GACATTAGCA TTCCAGTCTT ATGGAAAAAG GACGAGGGGA AGGAGAGGCT
5101   TTGCTCCTTC TATTAATCCC ATGAGCCAGG ACTTGCTTCT GTCACTTTTG
5151   TGATTCTTCC ACTTAACAGC ACCTGCTCAT GGGATGTCAT CCAGCATCAA
5201   GGAAAACTGG GATGTGGGTC CTTGTGCTGC TTGTACATTC TCAGAAAGGT
5251   TATGTGACCA AAAAAGGAAA TCTTGGGGCA ACCAGCAGTC TCTTCAGCCC
5301   CTGACTGTCT CTGATTCTGT GCTCACATCA AGATTTTTCA GGAACTCCTC
5351   AGAAATAATA AATGGTGGGG CAGAGAACAG AACTGGAGTC TCGTGCAGGA
5401   CTCCAGGGAC CAGGGGCTGG TATTGGACCT GCTCTTCATG TTGTGAACCA
5451   GGAAAACCCT TTAATTCTCT AGGCCTTAGC TTCATCTTAT GTTATATGAG
5501   GATAATACCA TAGACAGTCT TTAAAGAACA TCATAGCATG TTAAACAACA
5551   TGCTAAATGT TGGTGATACC ACAGTGAAAA AGACAGGCAT GACTTACTCC
5601   TTACGGATCT TCGGGTTTCA TGAGGAAGAC AAACATATCA TACCATACCT
5651   ATAGATGGAC AAACAGTTTA GTGCTCTGAG TGTGGATAAC AGAGGTTCTC
5701   CTTTTCCTCC CATTTCCTTT TTGGGCCAAT CAGAGCTGTG GCAGCTTGTC
5751   TCCCTAAGAG AGCTCATGAT GGATGCACTC ACTCCTGATG CTCCTCTATA
5801   CTCCCAGAGG AGGATGCATC TTCTTTCCAC CTGGAGAGCT CCTGCCCATG
5851   TGCATTCTTG GGATTCCAGA GCAAACGTGG CCTCTGATAG GCAAAAAAGA
5901   ACTCCTGAAT TTGTTCCTAA ATGGCACGCA CTCACCTCTA TTTTTCCCTT
5951   ATTTCATTTG CTTCTCATTC TCTATCTGGA GTTTGTTTAG GTTAATTTTT
6001   TTTTTCAGCC CACAATTTTG ACTGTCAACT TGGATTTAAC TTGAGAATCA
6051   CTCCTCTACT TTACCCCCCT CTAACATGTA TAATCGACAC ATAGTGGTGC
6101   TGGGTCCAAA GGGCTGGTGA AAAAATGGAT CATGAGTCAG CCCTGCTGGG
6151   CTCACATTCA TACTATATAA TATATAACCC CCCGGACAAA TAATAATCCTC
6201   TCTTTATACT CTAATTTCAT TATCTGCAAT ACAGGAATAA TACTAATTTT
6251   TACCTCCTAG GCTCTTCAGA TGATTAAAAG AGGCAATACC TAATAAAACTG
```

FIGURE 3, page 2 of 3

```
6301   TCAATCAGCT GCTGTTATTC TCCCAAATTA GACCTAATCC TCATTCTCCA
6351   GTTGAAATTT GCATGAATAT CTCTCTTTAC AACCCAAGCC CTACACTTCT
6401   CCTATTTCCA CTCATGGACT CCTCTCATAC AAATGTTTGC ATCAACAAAG
6451   AAACGCTACC AAAGATCTCC CGAAAGAGAG AATGAAATAG GTTTACATTG
6501   TGTATACTCA GCAGAACACT TAGTAGTCCC CCATACATAT TCCCACACTT
6551   CAATTACCTG CTGCAGTGGC ACTCAGGCTC ACCCTCACTT ACTCTTTCCT
6601   CTGTTCTATT GCTGAGCAAT TCAGCTCAGA CCCACACCCT ACCCAAACAC
6651   TGTGTACAAA ATGCTTCTAG GGGTTCGGCA AAGCCACACT GAGTCCTTAT
6701   TTTAAAGGCA CATCAGTGGT CAATTTCAGG TTTTGGGCAC TCATCAATCA
6751   TTCTTCTCAA CACAGATAGA GCTGTCCACA AATAGAATTC TGATGAATGA
6801   AATTTTCTTC ATCTAATTAT ATGTGTGTGT TCTAATGCCT TACATTGTGC
6851   TTTCATTTTT ATTTTCCATT TCATCCAAAT CTACCATTGC CATTAGGCTT
6901   CTCATGCATG CATTCCTTCA TTGAATGAAC GTTTATGAAA AGCACATTGT
6951   GCTGCTTATG GAATAGGCAC TAGGAGTATA AAATGTAAAA TGTGGTCCTG
7001   TCTGCAATGA CTGACACACT GAGTTATTTC TCACCCACCA GGTCCCGCCA
7051   TTTTCACACA TCCTAGCGAA GATCCCATTT TCCTCTGGTT CATAATGCAT
7101   GATCTTTTTT CCTGTCCAGA GATGACCAGT CCTGGTCATG AGGGTGTCAC
7151   AACCACCTCT TTGTGTATCT GAATTCCTCC ACCTGAGAGA AAATTTCAGG
7201   CCCAGGATAG AGTAATCATC GGGTCCACAG CACTGGCTAG ATGAGTGGGG
7251   GTGTTTTGAT CCTAATGTTA TCCCCATGTC AGCACAGAAC TTGTGTGGCA
7301   GTAGAGAGAG GTCAGGCTTC AGAGTCAACA AGAACTGGAT TTCAAACTGG
7351   ATTTGAGGAC CCCCACCTTT TGATAGGTGA CTTATTCTCT GCGAGTCTCT
7401   GATCTCTCCT CTTTAAATGA GGACAGTAAA TCCCACATGG CAGGGTGTCC
7451   GGGAGAATCA GAGATCAAAC AGCTGGTGAT CACATCTGGT TTCTGTTTCC
7501   AGGGTCATCA GACTGGGGTT TCTGAGCATG GATTCAACCA TCCCAGTCTT
7551   GGGTACAGAA CTGACACCAA TCAACGGACG TGAGGAGACT CCTTGCTACA
7601   AGCAGACCCT GAGCTTCACG GGGCTGACGT GCATCGTTTC CCTTGTCGCG
7651   CTGACAGGAA ACGCGGTTGT GCTCTGGCTC CTGGGCTGCC GCATGCGCAG
7701   GAACGCTGTC TCCATCTACA TCCTCAACCT GGTCGCGGCC GACTTCCTCT
7751   TCCTTAGCGG CCACATTATA TGTTCGCCGT TACGCCTCAT CAATATCCGC
7801   CATCCCATCT CCAAAATCCT CAGTCCTGTG ATGACCTTTC CCTACTTTAT
7851   AGGCCTAAGC ATGCTGAGCG CCATCAGCAC CGAGCGCTGC CTGTCCATCC
7901   TGTGGCCCAT CTGGTACCAC TGCCGCCGCC CCAGATACCT GTCATCGGTC
7951   ATGTGTGTCC TGCTCTGGGC CCTGTCCCTG CTGCGGAGTA TCCTGGAGTG
8001   GATGTTCTGT GACTTCCTGT TTAGTGGTGC TGATTCTGTT TGGTGTGAAA
8051   CGTCAGATTT CATTACAATC GCGTGGCTGG TTTTTTTATG TGTGGTTCTC
8101   TGTGGGTCCA GCCTGGTCCT GCTGGTCAGG ATTCTCTGTG GATCCCGGAA
8151   GATGCCGCTG ACCAGGCTGT ACGTGACCAT CCTCCTCACA GTGCTGGTCT
8201   TCCTCCTCTG TGGCCTGCCC TTTGGCATTC AGTGGGCCCT GTTTTCCAGG
8251   ATCCACCTGG ATTGGAAAGT CTTATTTTGT CATGTGCATC TAGTTTCCAT
8301   TTTCCTGTCC GCTCTTAACA GCAGTGGCCAA CCCCATCATT TACTTCTTCG
8351   TGGGCTCCTT TAGGCAGCGT CAAAATAGGC AGAACCTGAA GCTGGTTCTC
8401   CAGAGGGCTC TGCAGGACAC GCCTGAGGTG GATGAAGGTG GAGGGTGGCT
8451   TCCTCAGGAA ACCCTGGAGC TGTCGGGAAG CAGATTGGAG CAGTGAGGAA
8501   GAACCTCTGC CCTGTCAGAC AGGACTTTGA GAGCAATGCT GCCCTGCCAC
8551   CCTTGACAAT TATATGCATT TTTCTTAGCC TTCTGCCTCA GAAATGTCTC
8601   AGGGTCCCCA AGGCCCTTAC CA   (SEQ ID NO:3)
```

Features:
```
Start:       4300
Exon:        4300-4319
Intron:      4320-7502
Exon:        7503-8496
Stop:        8494
```

Chromosome Map Position:
Chromosome 3

FIGURE 3, page 3 of 3